AD-A037 678 STANFORD UNIV CALIF DEPT OF ELECTRICAL ENGINEERING F/G 12/1
LEVINSON- AND CHANDRASEKHAR-TYPE EQUATIONS FOR A GENERAL, DISCR--ETC(U)
DEC 76 B FRIEDLANDER, M MORF, T KAILATH F44620-74-C-0068
UNCLASSIFIED AFOSR-TR-77-0284 NL

1 OF 1
ADA037 678

END
DATE
FILMED
4-77

LEVINSON- AND CHANDRASEKHAR-TYPE EQUATIONS FOR A GENERAL DISCRETE-TIME LINEAR ESTIMATION PROBLEM

IEEE D & C Conference Dec. 1976

B. Friedlander, M. Morf, T. Kailath, L. Ljung†
Stanford Electronics Laboratories
Stanford University
Stanford, California 94305

Abstract

Recursive algorithms for the solution of linear least-squares estimation problems have been based mainly on state-space models. It has been known, however, that such algorithms exist for stationary time-series, using input-output descriptions (e.g., covariance matrices). We introduce a way of classifying stochastic processes in terms of their "distance" from stationarity that leads to a derivation of an efficient Levinson-type algorithm for arbitrary (nonstationary) processes. By adding structure to the covariance matrix, these general results specialize to state-space type estimation algorithms. In particular, the Chandrasekhar equations are shown to be the natural descendants of the Levinson algorithm.

1. Introduction

The problem of linear least squares estimation has been studied extensively and various methods of solution have been developed. These may be classified into estimation algorithms derived from input-output data or from other "external" system descriptions and algorithms derived from state-space or "internal" models. In the last decade the field of linear least-squares estimation has been dominated by state-estimation, in particular by the recursive Kalman-Bucy filter algorithm and its various versions, which rely heavily on the availability of state-space models.

In many applications, however, a state-space model is not readily available, and it would be preferable to have algorithms that use directly the covariance information of the observed process. The solution of the estimation problem is closely related to the problem of inverting the covariance matrix. Therefore, the computational efficiency of estimation algorithms is strongly dependent on the amount of computation required for inverting an appropriate matrix. For illustration we shall mention the important example of a stationary process and its Toeplitz-type covariance matrix. It has been shown [1-3] that by making use of the special structure of an $N \times N$ Toeplitz matrix, it can be inverted in $O(N^2)$ operations (multiplications and additions), compared to $O(N^3)$ operations required in general for the inversion of an arbitrary matrix. It has also been long known in certain fields (e.g., geophysical data processing [4] and speech compression studies [5]) that recursive solutions can be obtained for the prediction of stationary processes. In particular, the so-called Levinson algorithm computes the optimal-predictor in $O(N^2)$ operations.

It has seemed in the past that the prediction of nonstationary processes would require $O(N^3)$ operations unless we can impose a state-space structure on the signal and noise processes. However, it is not unreasonable to expect that between the highly structured Toeplitz matrix (or stationary process) and a completely arbitrary covariance matrix, there should exist matrices (or processes) with varying degrees of structure and that this structure could be somehow utilized in reducing the amount of computation involved in the estimation problem. That this is indeed possible has been first shown in [6] by introducing the concept of "shift (low) rank," see also [7-9], and subsequently in [10,11]. These results were motivated for discrete-time problems by the work of Levinson [14] and Golub [18], and for continuous-time by the Chandrasekhar-type equations and their further developments in [12,13].

In this paper we shall introduce α, an index of "distance from stationarity" of an arbitrary nonstationary process. We shall show how recursive solutions requiring of order αN^2 operations can be obtained for such processes with or without assuming a state-space structure. In the stationary case our solution reduces to known algorithms given in [14,15].

Finally, we shall show that, if the not necessarily stationary processes are known to come from state-space models, then this additional structural information can be used to reduce our general solution algorithm to the previously known Chandrasekhar-type equations. This means that we have been able to properly imbed the state-space assumption into a general input-output framework.

2. A General Linear Estimation Problem

We shall consider the problem of estimating a stochastic process $x(\cdot)$ from observations of a related process $y(\cdot)$. Let $y(\cdot)$ (p-dimensional) and $x(\cdot)$ (n-dimensional) have covariance matrices

---

†This work was supported by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF44-620-74-C-0068, and in part by the National Science Foundation under Contract NSF-Eng-75-18952 and the Joint Services Electronics Program under Contract N00014-75-C-0601.

3

$$R^N = \left[r_{i,j}\right] \qquad 0 \le i,j \le N \; ; \qquad r_{i,j} = E y_i y_j'$$

$$R^N_{xy} = \left[r^{xy}_{i,j}\right] \qquad 0 \le i,j \le N \; ; \qquad r^{xy}_{i,j} = E x_i y_j' \; .$$

The best linear least squares estimate of x_N given past $\{y_i,\ 0 \le i \le N-1\}$ has the form

$$\hat{x}(N|N-1) = \sum_{i=0}^{N-1} h_{xy}(N,i)\, y_i \; .$$

The optimal one-step ahead predictor $h_{xy}(N,\cdot)$ can be determined by using the well known orthogonality condition on the prediction error

$$x_N - \hat{x}(N|N-1) \perp y_k \; ; \qquad 0 \le k \le N-1$$

which means

$$0 = E\left[\left(x_N - \hat{x}(N|N-1)\right) y_k'\right] = r^{xy}_{N,k} - \sum_{i=0}^{N-1} h_{xy}(N,i)\, r_{i,k}$$

or in matrix form

$$h^N_{xy} R^{N-1} = \left[r^{xy}_{N,0}, \; \ldots, \; r^{xy}_{N,N-1}\right] , \tag{1}$$

where $h^N_{xy} \triangleq [h_{xy}(N,0), \ldots, h_{xy}(N,N-1)]$, an $n \times Np$ matrix.

Note that by setting $x = y$, we get an equation defining the predictor $h(\cdot,\cdot)$ of the observed process itself, i.e.,

$$\hat{y}(N|N-1) = \sum_{i=0}^{N-1} h(N,i)\, y_i \; ,$$

therefore

$$h^N R^{N-1} = \left[r_{N,0}, \; \ldots, \; r_{N,N-1}\right] . \tag{2}$$

This can be rewritten as

$$\left[-h^N, I\right] R^N = [0, \; \ldots, \; 0, E] \; , \tag{3}$$

where E is determined by the left-hand side of this equation. For estimating $x(\cdot)$ at a time instant within the observation interval $(0,N)$, we have to find the optimal filter (smoother) $H_{xy}(\cdot,\cdot;N)$, where

$$\hat{x}(k|N) = \sum_{i=0}^{N} H_{xy}(k,i;N)\, y_i \; .$$

Using the orthogonality condition again, now on the form

$$x_k - \hat{x}(k|N) \perp y_i \; , \qquad 0 \le i \le N$$

we get

$$\sum_{i=0}^{N} H_{xy}(k,i;N)\, r_{i,e} = r^{xy}_{k,e} \; , \qquad 0 \le k,e \le N$$

or in matrix form

$$H^N_{xy} R^N = R^N_{xy} \tag{4}$$

$$H^N_{xy} \triangleq \left[H_{xy}(i,j;N)\right] , \qquad 0 \le i,j \le N \; .$$

The last two matrix equations illustrate the fact that solving the estimation problem is closely related to the problem of inverting the covariance matrix, since for both solutions of the prediction- and the smoothing-problem we get

$$h^N_{xy} = \left[R^{N-1}\right]^{-1} \left[r^{xy}_{N,0}, \; \ldots, \; r^{xy}_{N,N-1}\right] \tag{5}$$

$$H^N_{xy} = R^N_{xy} \left[R^N\right]^{-1} . \tag{6}$$

3. Structural Assumptions

To aid in specifying the structural assumptions, we shall introduce some notation for the "shifted-difference" operators $\delta[\cdot]$ and $\rfloor[\cdot]$ that play a central role in this paper.

Let,

$$S \triangleq \left[s_{i,j}\right] , \qquad 0 \le i,j \le N \; ,$$

where $s_{i,j}$ are $p \times p$ matrices (i.e., S is a block matrix). Then define

$$\delta[S] \triangleq \begin{bmatrix} s_{1,1} & \cdots & s_{1,N} \\ \vdots & & \vdots \\ s_{N,1} & \cdots & s_{N,N} \end{bmatrix} - \begin{bmatrix} s_{0,0} & \cdots & s_{0,N-1} \\ \vdots & & \vdots \\ s_{N-1,0} & \cdots & s_{N-1,N-1} \end{bmatrix}$$

$$\rfloor[S] \triangleq \begin{bmatrix} s_{0,0} & \cdots & s_{0,N} \\ \vdots & & \vdots \\ s_{N,0} & \cdots & s_{N,N} \end{bmatrix} - \left[\begin{array}{c|ccc} 0 & & \cdots & 0 \\ \hline & s_{0,0} & \cdots & s_{0,N-1} \\ & \vdots & & \vdots \\ 0 & s_{N-1,0} & \cdots & s_{N-1,N-1} \end{array}\right]$$

$$= \left[\begin{array}{c|c} s_{0,0} & \cdots \; s_{0,N} \\ \hline \vdots & \delta[S] \\ s_{N,0} & \end{array}\right] .$$

We can now define the (block) displacement rank α of matrix S as

$$\alpha \triangleq \lceil \text{rank } \delta[S]/p \rceil \; ,$$

where

$$\lceil x \rceil \triangleq \text{smallest integer } m, \text{ such that } m \ge x \; .$$

Remarks

1. If $p = 1$ (i.e., when S has scalar entries) then
$$\alpha = \text{rank } \delta[S] \ .$$
2. When S is a Toeplitz matrix, then $\delta[S] = 0$ and therefore $\alpha = 0$. If S is an arbitrary matrix, $\delta[S]$ may be full rank and then $\alpha = N$. Therefore, α is bounded by $0 \le \alpha \le N$, and the actual value of α is an index of "distance from stationarity" of the matrix S.

Examples

Let T be a full Toeplitz matrix and $L_i(U_i)$ be lower (upper) Toeplitz matrices.

(i) $\quad S = LU \ , \quad \alpha \le 1 \ .$

It can be shown (check for a 3×3 case) that,
$$\lrcorner[LU] = lu'$$
where l is the first column of L and u' is the first row of U. Therefore, rank $\lrcorner[LU]$ is at most p.

(ii) $\quad S = T \ , \quad \alpha \le 2 \ .$

For Toeplitz matrices $\delta[T] = 0$, and in $\lrcorner[T]$ only the first row and column may be nonzero.

(iii) $\quad S = U \cdot L \ , \quad \alpha \le 3 \ .$

(iv) $\quad S = T_1 T_2 \ , \quad \alpha \le 4 \ .$

(v) $\quad S = \sum_{i=1}^{k} L_i U_i \ , \quad \alpha \le k \ .$

Since the matrix $\delta[S]$ has a rank $\le \alpha \cdot p$ (by the definition of α), we can always factor it (nonuniquely) as
$$\delta[S] = P_1 \, \Sigma \, P_2'$$
with a signature matrix Σ and P_1, P_2 being $Np \times \alpha p$ (block) matrices.

If S is a symmetric matrix, so is $\delta[S]$. In this case, a (nonunique) symmetric decomposition can always be found of the form
$$\delta[S] = P \, \Sigma \, P' \qquad (\text{i.e., } P_1 = P_2) \tag{7}$$
where Σ is again the signature matrix.

The question of how to find this decomposition and further details about the displacement rank α can be found in [10,11]. For our present discussion it suffices to know that given an arbitrary covariance matrix R of an observed process $y(\cdot)$, we can associate with it a number α and an $Np \times \alpha p$ matrix P such that $\delta[R] = P \Sigma P'$. An important example that further illustrates the meaning of α and P will be presented in Sec. 5.

We can now proceed in solving a general estimation problem by stating our assumptions on the crosscovariance R_{xy},
$$\delta[R_{xy}] = P_{xy} \, \Sigma \, P' \tag{8}$$
where P_{xy} is a $Nn \times \alpha p$ matrix.

Note that if the processes $x(\cdot)$ and $y(\cdot)$ are jointly stationary, then
$$\delta[R_{xy}] = 0 \ , \quad \delta[R] = 0 \ , \quad P_{xy} = 0 \ , \quad P = 0 \ .$$

The motivation behind this assumption is that in many problems the signal $x(\cdot)$ and the observations $y(\cdot)$ are connected by a linear relation of the form
$$y_i = Hx_i + v_i = z_i + v_i \tag{9}$$
where $v(\cdot)$ is white noise with unit intensity, uncorrelated with $x(\cdot)$. In this case,
$$\begin{aligned} r_{i,j} &= Ey_i y_j' = Ez_i z_j' + I \cdot \delta_{i,j} = Ez_i y_j' + I \cdot \delta_{i,j} \\ &= HEx_i y_j + I \cdot \delta_{i,j} = Hr^{xy}_{i,j} + I \cdot \delta_{i,j} \end{aligned} \tag{10}$$
$$\delta_{i,j} \triangleq \begin{cases} 0 & i \ne j \\ 1 & i = j \end{cases}$$
and R_{xy} will indeed satisfy the assumption (8). As a matter of fact,
$$R = \text{diag } \{H\} \, R_{xy} + I$$
and by operating with $\delta[\cdot]$ on both sides we get
$$P = \text{diag } \{H\} \, P_{xy} \ .$$

4. The Levinson-Type Algorithm for the Joint $\{x,y\}$ Process

Using the assumptions stated in the previous section on the covariance information, we can now give a set of recursions for computing h^N_{xy}.

$$h^{m+1}_{xy} = \left[0, h^m_{xy}\right] + E_m M_m^{-1} B^{m'} \ , \qquad h^1_{xy} = r^{xy}_{1,0} r^{-1}_{0,0} \tag{11a}$$

$$A^{m+1} = \begin{bmatrix} 0 \\ A^m \end{bmatrix} - \begin{bmatrix} B^m \\ 0 \end{bmatrix} M_m^{-1} C_m' \ , \qquad A^0 = I \tag{11b}$$

$$B^{m+1} = \begin{bmatrix} B^m \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ A^m \end{bmatrix} N_m^{-1} C_m \ , \qquad B^0 = [I, 0, \ldots, 0] \tag{11c}$$

$$N_{m+1} = N_m - C_m M_m^{-1} C_m' \ , \qquad N_0 = r_{0,0} \tag{12a}$$

$$M_{m+1} = M_m - C_m' N_m^{-1} C_m \ , \qquad M_0 = \begin{bmatrix} r_{0,0} & 0 \\ 0 & -\Sigma \end{bmatrix} \tag{12b}$$

Define $\tilde{\alpha} = (\alpha+1)p$ and $\tilde{m} = mp$ (the size of R^m). The dimensions of h_{xy}, A^m, B^m, N_m, M_m, E_m, C_m are $n\times\tilde{m}$, $(\tilde{m}+p)\times p$, $(\tilde{m}+p)\times\tilde{\alpha}$, $p\times p$, $\tilde{\alpha}\times\tilde{\alpha}$, $n\times\tilde{\alpha}$, and $p\times\tilde{\alpha}$, respectively. The quantities E_m, C_m have to be computed at each step by

$$E_m = \begin{bmatrix} 0 & h^m_{xy} \end{bmatrix} \begin{bmatrix} r_{0,0} & 0 \\ \vdots & \\ r_{m,0} & P^{m-1}\Sigma \end{bmatrix} + \begin{bmatrix} r^{xy}_{m+1,0}, p^{xy}_m\Sigma \end{bmatrix} \tag{13}$$

$$C_m = \begin{bmatrix} 0 & A'_m \end{bmatrix} \begin{bmatrix} r_{0,0} & 0 \\ \vdots & \\ r_{m+1,0} & P^m\Sigma \end{bmatrix}$$

$$= \begin{bmatrix} r_{m+1,0}, \ldots, r_{m+1,m} \end{bmatrix} B^m - \begin{bmatrix} 0, p_m \end{bmatrix} M_m \tag{14}$$

where p_m, p^{xy}_m are the m^{th} block row of P, P^{xy}, respectively, and

$$P^m \triangleq \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix}.$$

By counting the number of operations required at the m^{th} step of the recursion, we get (assuming $p \ll \tilde{m}$ and ignoring terms accordingly) $(2n+3p)\tilde{m}\tilde{\alpha}$ multiplications. Finding h^m_{xy} will therefore require $\sim(1.5+n/p)\tilde{m}^2\tilde{\alpha}$ multiplications. The proof of the recursions above is somewhat lengthy and is given in Appendix A. It is however, important to note that the auxiliary quantities A^m, B^m also obey the following equations.

$$R^m A^m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ N_m \end{bmatrix} \tag{15}$$

$$R^m B^m M_m^{-1} = \begin{bmatrix} I & 0 \\ 0 & \\ \vdots & P^{m-1} \\ 0 & \end{bmatrix}. \tag{16}$$

The first equation implies that

$$A^m = \begin{bmatrix} -h^m, I \end{bmatrix}'$$

so that A^m is just the optimal *predictor defined* in the previous section. Note also that in the stationary case

$$R^m B^m = \begin{bmatrix} M_m \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

in which case B^m is the smoothing filter for estimating y_0 given $\{y_i, 1 \le i \le N\}$, or in this case also the so-called "backward predictor" [19, 20].

Recursive solutions of this type were developed for the stationary case by Levinson [14] for computing $h(\cdot,\cdot)$ and by Wiggins and Robinson [15] for computing $h_{xy}(\cdot,\cdot)$. Indeed, when we take $\alpha = 0$, $P = 0$ the equations (11b), (11c), (12) reduce to the Levinson algorithm and (11), (12) can be shown to be equivalent to the equations of Wiggins and Robinson. Thus, the stationary case is nicely imbedded in our framework.

5. State Space Structure and Chandrasekhar-Type Equations

The results described so far are quite general and do not require a state space structure. We shall now show how by imposing more structure on the covariance matrices, the Chandrasekhar-type equations can be derived from the Levinson-type equation presented in the previous section.

Let $y(\cdot)$ and $x(\cdot)$ be the output and the state vectors of a linear system driven by white noise, i.e.,

$$\begin{aligned} x_{i+1} &= F_i x_i + u_i \\ y_i &= H_i x_i + v_i \end{aligned} \tag{17}$$

$$Eu_i u'_j = Q_i \delta_{i,j}, \qquad Ev_i v'_j = I\delta_{i,j}$$

$$Eu_i v_j = Ex_0 u'_j = 0 .$$

In this case,

$$Ex_{i+1} y'_j = F_i (Ex_i y'_j) + Eu_i y'_j ,$$

where the last term is zero for $i \ge j$. Therefore,

$$r^{xy}_{i+1,j} = F_i r^{xy}_{i,j}, \qquad i \ge j . \tag{18}$$

Also, as already noted earlier,

$$r_{i,j} = H_i r^{xy}_{i,j} + I\delta_{i,j} . \tag{19}$$

In the following discussion, we shall therefore assume that R, R_{xy} obey assumptions (18), (19) which are somewhat weaker than the state-space assumption above.

The optimal filter $h_{xy}(\cdot,\cdot)$ was shown to obey equation (1). Therefore,

$$\sum_{i=0}^{t} h_{xy}(t+1,i)\, r_{i,s} = r^{xy}_{t+1,s}$$

$$\sum_{i=0}^{t-1} h_{xy}(t,i)\, r_{i,s} = r^{xy}_{t,s} .$$

Subtracting these last equations and using (18), we get

$$\sum_{i=0}^{t-1}\left[h_{xy}(t+1,i) - h_{xy}(t,i)\right] r_{i,s} + h_{xy}(t+1,t)\, r_{t,s}$$

$$= r^{xy}_{t+1,s} - r^{xy}_{t,s} = (F_t - I)\, r^{xy}_{t,s}$$

or

$$\sum_{i=0}^{t-1}\left[h_{xy}(t+1,i) - h_{xy}(t,i)\right] r_{i,s}$$

$$= \left(F_t - I - h_{xy}(t+1,t)\, H_t\right) r^{xy}_{t,s}$$

and by comparison with (1)

$$h_{xy}(t+1,i) - h_{xy}(t,i)$$

$$= \left(F_t - I - h_{xy}(t+1,t)\, H_t\right) h_{xy}(t,i)$$

or

$$h_{xy}(t+1,i) = \left(F_t - h_{xy}(t+1,t)\, H_t\right) h_{xy}(t,i) \, . \tag{20}$$

Using the estimator equation

$$\hat{x}(t|t-1) = \sum_{i=0}^{t-1} h_{xy}(t,i)\, y_i$$

we get a recursive formula for the estimate,

$$\hat{x}(t+1|t) = \sum_{i=0}^{t} h_{xy}(t+1,i)\, y_i$$

$$= \left(F_t - h_{xy}(t+1,t)\, H_t\right) \sum_{0}^{t-1} h_{xy}(t,i)\, y_i$$

$$+ h_{xy}(t+1,t)\, y_t$$

$$= F_t\hat{x}(t|t-1) + h_{xy}(t+1,t)$$

$$\cdot \left(y_t - h_{xy}(t+1,t)\, H_t\hat{x}(t|t-1)\right) , \tag{21}$$

the usual Kalman-filter equation for the state estimates. Note that only $h_{xy}(t+1,t)$ is required under the state-space structure assumptions!

In the following, we shall assume that F and H are constant. By considering the defining equations (1), (2) of $h(\cdot,\cdot)$ and $h_{xy}(\cdot,\cdot)$, it is easy to see that assumption (19) leads to

$$h(t,s) = Hh_{xy}(t,s) \, .$$

Therefore, from the definitions (13), (14) of E_m and C_m it follows that

$$C_m = HE_m \, , \tag{22}$$

and from the Levinson recursions (11a), (11c) for h_{xy} and B we have

$$h_{xy}(t+1,t) = h_{xy}(t,t-1) + E_t M_t^{-1}\left(B^t_t\right)'$$

$$B^t_t = -N^{-1}_{t-1}C_{t-1} \, , \qquad B^t_t = \text{the last block row of } B^t \, ,$$

so that

$$h_{xy}(t+1,t) = h_{xy}(t,t-1) - E_t M_t^{-1} E'_{t-1} H' N^{-1}_{t-1} \, . \tag{23}$$

This recursion can be rewritten in another form that is easier to compare with the usual Chandrasekhar equations,

$$h_{xy}(t+1,t)\, N_t = h_{xy}(t,t-1)\, N_{t-1} - FE_{t+1}M^{-1}_{t+1}E'_{t-1}H' \tag{24}$$

The necessary algebra to derive this from (23) is given in Appendix B.

In Appendix C we shall also prove that

$$E_t = \left(F - h_{xy}(t,t-1)\, H\right) E_{t-1} \, . \tag{25}$$

Finally, note that (12), (24), (25) provide a complete set of recursions for computing $h_{xy}(t+1,t)$, which are of the Chandrasekhar-type. Indeed, a comparison of our results to those presented in [16] shows that equations (24), (25), (12a), (12b) are precisely equations (16), (13'), (15), (14) of [16] if the following change of notation is made.

$$h_{xy}(t+1,t) \leftrightarrow K_g(t) \quad \left(\text{or} \quad h_{xy}(t+1,t)\, N_t \leftrightarrow K_t\right) ,$$

$$N_t \leftrightarrow R^r(t) \, , \qquad M_t \leftrightarrow -M_t \, , \qquad E_t \leftrightarrow y_t \, .$$

We see therefore that the Chandrasekhar equations are naturally induced by the Levinson-type recursions when the covariance matrices have a special structure. Furthermore, the parameter α_0 which appears in the Chandrasekhar algorithm can now be shown to have a meaning in terms of α (or $\tilde{\alpha}$), the displacement rank of R (see also [8,9]).

To see this, let us recall that for a time-invariant state-space model

$$r_{i,j} = \begin{cases} HF^{i-j}\pi_j H' & i > j \\ H\pi_i H' + I & i = j \\ H\pi_i (F^{j-i})' H' & i < j \end{cases} \tag{26}$$

where

$$\pi_i = Ex_i x_i'$$

and

$$\pi_{i+1} = F\pi_i F' + Q \, . \tag{27}$$

so that,

$$r_{i+1,j+1} - r_{i,j} = HF^{i-j}(\pi_{j+1} - \pi_j)\, H'$$

$$= HF^i(\pi_1 - \pi_0)(F^j)'\, H' \, .$$

Writing this in matrix form gives

$$\delta R^N = O\ (\pi_1 - \pi_0)\ O' \tag{28}$$

where

$$O \triangleq \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{N-1} \end{bmatrix}, \quad \text{an extended observability matrix.}$$

Let us take for simplicity the scalar case (i.e., $p = 1$, $r_{i,j}$ scalar). Assuming that the system is observable, O will be a full rank matrix, and therefore

$$\alpha = \text{rank } \delta R^N = \text{rank } (\pi_1 - \pi_0) .$$

Now the Chandrasekhar equations involve a parameter α_0 defined as

$$\alpha_0 = \text{rank } \left[(\pi_1 - \pi_0) - F\pi_0 H' (H\pi_0 H' + I)^{-1} H\pi_0 F' \right]$$

and since the second term in the brackets is of rank 1,

$$\alpha - 1 \leq \alpha_0 \leq \alpha + 1 .$$

6. Conclusions

We have shown how recursive solutions can be obtained for the optimal predictor with covariance (or input-output) data, whether or not state-space models are available. The complexity of these algorithms depends on a measure α of the "distance" from stationarity of the signal and observed processes.

A similar approach makes it possible to derive a recursive solution for the optimal (filter) smoother H_{xy}. The details will not be presented here (see [10] for a partial treatment), but it is important to note that H_{xy} can be computed in $O(\alpha N^2)$ operations, instead of $O(N^3)$ required for a direct solution of equation (4).

In the special case where the processes are known to come from a constant-parameter state-space model, the distance from stationarity α coincides with a parameter describing the computational reductions obtainable by using the previously known [16] Chandrasekhar equations. Moreover, these general recursions reduce naturally to the Chandrasekhar equations in this and actually also in some more general cases. Note, for example, that we made no assumption on Q and our derivation holds when it is time varying. Note also that our approach leads to a derivation of the Chandrasekhar equations that does not mention the Riccati equations, which was at the heart of the original derivation [16]. Actually, the general time-varying Riccati equation can also be imbedded in the framework presented here (see [17]).

Finally, we should note that these discrete-time results have close continuous-time analogs presented in [13] for the general problem of solving some integral equations, and in [12] dealing specifically with the estimation problem. In fact, it was these results that provided the immediate motivation for the discrete-time analysis presented here.

Appendix A

Proof of the Levinson-Type Algorithm for the Joint (x,y) Process

The defining equation for h^m_{xy} was given as

$$h^m_{xy} R^{m-1} = \left[r^{xy}_{m,0} \ \cdots \ r^{xy}_{m,m-1} \right] \tag{A1}$$

Therefore,

$$\begin{aligned}
&\left[0 \ h^m_{xy}\right] R^m \\
&= \left[0 \ h^m_{xy}\right] \left\{ \begin{bmatrix} 0 \ \ldots \ 0 \ \ldots \ 0 \\ \vdots \quad R^{m-1} \\ 0 \end{bmatrix} + \begin{bmatrix} r_{0,0} \ \cdots\cdots \ r_{0,m} \\ \vdots \quad P^{m-1} \Sigma P^{m-1} \\ r_{m,0} \end{bmatrix} \right\} \\
&= \left[0, r^{xy}_{m,0} \ \cdots \ r^{xy}_{m,m-1}\right] + \left[\Delta_1, 0 \ \ldots \ 0\right] \\
&\quad + \left(h^m_{xy} P^{m-1} \Sigma\right) \left[0, P^{m-1}\right] \\
&= \left[r^{xy}_{m+1,0} \ \cdots \ r^{xy}_{m+1,m}\right] - \left[r^{xy}_{m+1}, p^{xy}_m \Sigma P^{m-1}\right] \\
&\quad + \left[\Delta_1, 0 \ \ldots \ 0\right] + \left(h^m_{xy} P^{m-1} \Sigma\right) \left[0, P^{m-1}\right]
\end{aligned}$$

where

$$\Delta_1 = \left[0 \ h^m_{xy}\right] \begin{bmatrix} r_{0,0} \\ \vdots \\ r_{m,0} \end{bmatrix}$$

This can be rewritten as

$$\left[0 \ h^m_{xy}\right] R^m = h^{m+1}_{xy} R^m - E_m \begin{bmatrix} I & 0 \ \ldots \ 0 \\ 0 & P^{m-1'} \end{bmatrix} \tag{A2}$$

where

$$E_m \triangleq -\left[0 \ h^m_{xy}\right] \begin{bmatrix} r_{0,0} & 0 \\ \vdots & P^{m-1}\Sigma \\ r_{m,0} & \end{bmatrix} + \left[r^{xy}_{m+1}, p^{xy}_m \Sigma\right] .$$

Define auxiliary quantities A^m, B^m as

$$R^m A^m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ N_m \end{bmatrix} \tag{A3}$$

where the last block row of A^m is the identity matrix.

$$R^m B^m M_m^{-1} = \begin{bmatrix} I & 0 \\ 0 & \\ \vdots & P^{m-1} \\ 0 & \end{bmatrix} \qquad (A4)$$

Note that from (A2) and (A4) it follows that

$$h_{xy}^{m+1} = \begin{bmatrix} 0 & h_{xy}^m \end{bmatrix} + E_m M_m^{-1} B^{m'} \qquad (A5)$$

and from the defining equation, $h_{xy}^1 = r_{1,0}^{xy} r_{0,0}^{-1}$. Using a similar approach, the recursions for A^m, B^m are derived.

$$R^{m+1}\begin{bmatrix} 0 \\ A^m \end{bmatrix} = \left\{ \begin{bmatrix} r_{0,0} & \cdots & r_{0,m+1} \\ \vdots & & \\ r_{m+1} & & R^m \end{bmatrix} + \begin{bmatrix} 0 & \cdots\cdots & 0 \\ \vdots & P^m \Sigma P^{m'} & \\ 0 & & \end{bmatrix} \right\} \begin{bmatrix} 0 \\ A^m \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ N_m \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \\ \vdots & P^m \\ 0 & \end{bmatrix} F_m , \qquad (A6)$$

where

$$F_m \triangleq \begin{bmatrix} 0 & A_m' \end{bmatrix} \begin{bmatrix} r_{0,0} & 0 \\ \vdots & \\ r_{m+1,0} & P^m \Sigma \end{bmatrix}$$

$$R^{m+1}\begin{bmatrix} B^m \\ \\ 0 \end{bmatrix} M_m^{-1} = \begin{bmatrix} I & 0 \\ 0 & P^{m-1} \\ \vdots & \\ 0 & - - - -1 \\ k_m & M_m \end{bmatrix} \qquad (A7)$$

$$k_m \triangleq \begin{bmatrix} r_{m+1,0} & \cdots & r_{m+1,m} \end{bmatrix} B^m$$

Combining the defining equations of A^m, B^m and (A6), (A7) gives

$$A^{m+1} = \begin{bmatrix} 0 \\ A^m \end{bmatrix} - \begin{bmatrix} B^m \\ 0 \end{bmatrix} M_m^{-1} F_m' , \qquad A_0 = I . \qquad (A8)$$

$$N_{m+1} = N_m - \left(k_m - [0, p_m]\ M_m\right) M_m^{-1} F_m'$$

$$= N_m - C_m M_m^{-1} F_m' , \qquad N_0 = r_{0,0} . \qquad (A9)$$

where

$$C_m \triangleq k_m - [0, p_m]\ M_m .$$

Also

$$B^{m+1} = \begin{bmatrix} B^m \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ A^m \end{bmatrix} N_m^{-1} C_m , \qquad B_0 = [I \quad 0 \ \ldots \ 0] . \qquad (A10)$$

and

$$M_{m+1} = M_m - F_m' N_m^{-1} C_m , \qquad M_0 = \begin{bmatrix} r_{0,0} & 0 \\ 0 & \epsilon \end{bmatrix} . \qquad (A11)$$

To verify (A8), (A9), we premultiply the equations by R^{m+1} and check that the right-hand side of these equations satisfy the defining equations (A3), (A4).

Finally, it remains to show that $C_m = F_m$. The proof is lengthy and shall be omitted. It can be found in [10,11].

Appendix B

An Alternative form of the Recursion for $h_{xy}(t+1\,t)$

$$h_{xy}(t+1,t)\ N_t = \left(h_{xy}(t,t-1) - E_t M_t^{-1} E_{t-1}' H' N_{t-1}^{-1}\right)$$

$$\cdot \left(N_{t-1} - HE_{t-1} M_{t-1}^{-1} E_{t-1}' H\right)$$

$$= h_{xy}(t,t-1)\ N_{t-1} - E_t M_t^{-1} E_{t-1}' H'$$

$$- h_{xy}(t,t-1)\ HE_{t-1} M_{t-1}^{-1} E_{t-1}' H'$$

$$+ E_t M_t^{-1} E_{t-1}' H' N_{t-1}' HE_{t-1} M_{t-1}^{-1} E_{t-1}' H$$

$$= h_{xy}(t,t-1)\ N_{t-1}$$

$$+ E_t M_t^{-1} \left(E_{t-1}' H' N_{t-1}^{-1} HE_{t-1} + -M_{t-1}\right)$$

$$M_{t-1}^{-1} E_{t-1}' H$$

$$- h_{xy}(t,t-1)\ HE_{t-1} M_{t-1}^{-1} E_{t-1}' H'$$

$$= h_{xy}(t,t-1)\ N_{t-1}$$

$$- \left(E_t + h_{xy}(t,t-1) HE_{t-1}\right) M_{t-1}^{-1} E_{t-1}' H'$$

In Appendix C, we shall show that

$$E_t = \left(F - h_{xy}(t,t-1)\ H\right) E_{t-1}$$

therefore

$$h_{xy}(t+1,t)\ N_t = h_{xy}(t,t-1)\ N_{t-1} - FE_{t-1} M_{t-1}^{-1} E_{t-1}' H' .$$

Appendix C

The Recursion of E_t

As a first step, we shall show that

$$h_{xy}(t,s) = \Big(F - h_{xy}(t,t-1)\,H\Big)\, h_{xy}(t-1,s)$$

$$0 \le s \le t-2 \qquad \text{(C1)}$$

From the defining equation (1) for h^t_{xy}, we know that

$$\left(h^t_{xy} - \left[h^{t-1}_{xy}, 0\right]\right) R^{t-1}$$

$$= \left[r^{xy}_{t,0}, \ldots, r^{xy}_{t,t-1}\right] - \left[r^{xy}_{t-1,0} \ldots r^{xy}_{t-1,t-2}, \Delta_1\right]$$

where

$$\Delta_1 = \left[h^{t-1}_{xy}, 0\right]\begin{bmatrix} r_{0,t-1} \\ \\ r_{t-1,t-1}\end{bmatrix}.$$

Using

$$r^{xy}_{t,s} = F r^{xy}_{t-1,s}$$

gives

$$\left(h^t_{xy} - \left[h^{t-1}_{xy}, 0\right]\right) R^{t-1}$$

$$= (F - I)\left[r^{xy}_{t-1,0}, \ldots, r^{xy}_{t-1,t-2}, \Delta_1\right]$$

$$\left[0, \ldots, 0, \Delta_2\right]$$

where Δ_2 is whatever is necessary to satisfy the last equation. Hence,

$$h^t_{xy} - \left[h^{t-1}_{xy}, 0\right] = (F - I)\left[h^t_{xy}, 0\right]$$

$$+ h_{xy}(t,t-1)\left[-h^{t-1} I\right] \qquad \text{(C2)}$$

Using $h(t,s) = Hh_{xy}(t,s)$, we can rewrite (C2) as

$$h^t_{xy} = \Big(F - h_{xy}(t,t-1)\,H\Big)\left[h^{t-1}_{xy} . \; 0\right]$$

$$+ \left[0 \ldots 0, h_{xy}(t,t-1)\right]$$

which proves (C1).

Let us rewrite the definition (13) of E_t as

$$E_t = -\left[0, h^t_{xy}\right]\begin{bmatrix} r_{0,0} & 0 \\ \vdots & P^{t-2}\Sigma \\ r_{t-1,0} & \\ 0 & 0 \end{bmatrix}$$

$$- h_{xy}(t,t-1)\left[r_{t,0} \;\vdots\; P_{t-1}\Sigma\right] + \left[r^{xy}_{t+1,0} \;\vdots\; P^{xy}_t\Sigma\right].$$

$$\text{(C3)}$$

It is also true that

$$\left[r_{t,0} \;\vdots\; P_{t-1}\Sigma\right] = H\left[r^{xy}_{t,0} \;\vdots\; P^{xy}_t\Sigma\right] \qquad \text{(C4)}$$

$$\left[r^{xy}_{t+1,0} \;\vdots\; P^{xy}_t\Sigma\right] = F\left[r^{xy}_{t,0} \;\vdots\; P^{xy}_{t-1}\Sigma\right] \qquad \text{(C5)}$$

Using (C1), (C4), and (C5), we can rewrite (C3) as

$$E_t = \Big[F - h_{xy}(t,t-1)\,H\Big]\, h^{t-1}_{xy}\begin{bmatrix} r_{1,0} & \\ \vdots & P^{t-2}\Sigma \\ r_{t-1,0} & \end{bmatrix}$$

$$+ \Big(F - h_{xy}(t,t-1)\,H\Big)\left[r^{xy}_{t,0} \;\vdots\; P^{xy}_{t-1}\Sigma\right]$$

$$= \Big(F - h_{xy}(t,t-1)\,H\Big)\, E_{t-1}\,. \qquad \text{(C6)}$$

REFERENCES

1. W. Trench, "An Algorithm for the Inversion of Finite Hankel Matrices," SIAM J. Appl. Math., vol. 13, pp. 1102-1107, 1965.

2. S. Zohar, "The Solution of a Toeplitz Set of Linear Equations," J. Assoc. Comput. Math., vol. 21, pp. 272-276, 1974.

3. H. Akaike, "Bloçk Toeplitz Matrix Inversion," SIAM J. Appl. Math., vol. 24, no. 2, March 1975.

4. L. C. Wood and S. Treitel, "Seismic Signal Processing," Proc. IEEE, vol. 63, no. 4, pp. 649-661, April 1975.

5. J. Makhoul, "Linear Prediction: A Tutorial Review," Proc. IEEE, vol. 63, no. 4, pp. 561-580, April 1975.

6. M. Morf, "Fast Algorithms for Multivariable Systems," Ph.D. Dissertation, Dept. of Elec. Engr., Stanford University, Stanford, Ca., 1974.

7. G. S. Sidhu, T. Kailath and M. Morf, "Development of Fast Algorithms via Innovations Decomposition," Proc. of the Seventh Hawaii Intl. Conf. on Systems Sciences, pp. 192-195, Honolulu, Hawaii, January 1974.

8. M. Morf, T. Kailath and B. Dickinson, "General Speech Models and Linear Estimation Theory," reprinted from Speech Recognition, New York: Academic Press, 1975.

9. B. W. Dickinson, T. Kailath and M. Morf, "Canonical Matrix Fraction and State-Space Descriptions for Deterministic and Stochastic Linear Systems," IEEE Trans. on Automatic Control, vol. AC-19, no. 6, pp. 656-666, December 1974.

10. B. Friedlander, "Scattering Theory and Linear Least-Squares Estimation," Ph.D. Dissertation, Dept. of Elec. Engr., Stanford University,

Stanford, Ca., 1974.

11. B. Friedlander, M. Morf, T. Kailath and L. Ljung, "New Inversion Formulas for Matrices Classified in Terms of Their Distance From Toeplitz Matrices," submitted to SIAM J. Appl. Math.

12. T. Kailath, L. Ljung and M. Morf, "Recursive Input-Output and State-Space Solutions for Continuous-Time Linear Estimation Problems," submitted for publication.

13. T. Kailath, L. Ljung and M. Morf, "A New Approach to the Determination of Fredholm Resolvents of Nondisplacement Kernels," special issue in honor of M. G. Krein of Advances in Mathematics, 1977.

14. N. Levinson, "The Wiener RMS (Root-Mean Square) Error Criterion in Filter Design and Prediction," J. Math. Phys., vol. 25, pp. 261-278, January 1947.

15. R. A. Wiggins and E. A. Robinson, "Recursive Solution to the Multichannel Filtering Problem," J. Geophysical Res., vol. 70, no. 8, pp. 1885-1891, 1965.

16. M. Morf, G. S. Sidhu and T. Kailath, "Some New Algorithms for Recursive Estimation in Constant, Linear, Discrete-Time Systems," IEEE Trans. on Automatic Control, vol. AC-19, no. 4, pp. 315-328, August 1974.

17. B. Friedlander, T. Kailath and M. Morf, "Imbedding the Time-Varying Riccati Equation into Recursive Input-Output Equations," submitted for publication.

18. P. E. Gill, G. H. Golub, W. Murray and M. A. Saunders, "Methods for Modifying Matrix Factorization," Math. Comp., vol. 28, pp. 505-525, 1974.

19. T. Kailath, "A View of Three Decades of Linear Filtering Theory," IEEE Trans. on Information Theory, vol. IT-20, no. 2, pp. 165-181, March 1974.

20. P. Whittle, Prediction and Regulation, New York: Van Nostrand Reinhold, 1963.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR - TR - 77 - 0284	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) LEVINSON-AND CHANDRASEKHAR-TYPE EQUATIONS FOR A GENERAL, DISCRETE-TIME LINEAR ESTIMATION PROBLEM		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) B. Friedlander, M. Morf, T. Kailath and L. Ljung		8. CONTRACT OR GRANT NUMBER(s) F44620-74-C-0068 NSF-Eng-75-18952
9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford University Dept of Electrical Engineering Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, WAshington, DC 20332		12. REPORT DATE Dec 76
		13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 13 p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Levinson-and Chandrasekhar-type equations
discrete-time linear estimation
state-space models
stationary time-series
covariance matrices

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Recursive algorithms for the solution of linear least-squares estimation problems have been based mainly on state-space models. It has been known, however, that such algorithms exist for stationary time-series, using input-output descriptions (e.g., covariance matrices). We introduce a way of classifying stochastic processes in terms of their "distance" from stationarity that leads to a derivation of an efficient Levinson-type algorithm for arbitrary (nonstationary) processes. By adding structure to the covariance matrix, these general results specialize to state-space type estimation next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 400 852

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20 Abstract

cont

algorithms. In particular, the Chandrasekhar equations are shown to be the natural descondants of the Levinson algorithm.